JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2008, **89,** 183–207 NUMBER 2 (MARCH)

DETERMINANTS OF IMITATION OF HAND-TO-BODY GESTURES IN 2- AND 3-YEAR-OLD CHILDREN

MIHELA ERJAVEC AND PAULINE J. HORNE

BANGOR UNIVERSITY

Twenty children, ten 2-year-olds and ten 3-year-olds, participated in an AB procedure. In the baseline phase, each child was trained the same four matching relations to criterion under intermittent reinforcement. During the subsequent imitation test, the experimenter modeled a total of 20 target gestures (six trials each) interspersed with intermittently reinforced baseline trials. In each session, target gestures were selected in a pre-randomized sequence from: Set 1—ear touches; Set 2—shoulder touches; Set 3—midarm touches; and Set 4—wrist touches; subjects' responses to targets were not reinforced. In each target set, half the gestures featured in nursery matching games and were termed *common targets* whereas the remainder, which were topographically similar but did not feature in the games, served as *uncommon targets.* The children produced significantly more matching responses to common target models than to uncommon ones. Common responses were also produced as mismatches to uncommon target models more often than vice versa. Response accuracy did not improve over trials, suggesting that "parity" did not serve as a conditioned reinforcer. All children showed a strong bias for "mirroring"—responding in the same hemispace as the modeler. The 2-year-olds produced more matching errors than the 3-year-olds and most children showed a bias for responding with their right hands. The strong effects of training environment (nursery matching games) are consistent with a Skinnerian account, but not a cognitive goal theory account, of imitation in young children.

Key words: imitation, generalized imitation, trained matching, manual gestures, matching games, naming, young children

---

Many psychologists, of various theoretical persuasions, take the view that imitation is established early in human development and that, in turn, it facilitates the acquisition of other important behavioral repertoires exclusive to humans, such as verbal behavior (Hurley & Chater, 2005, pp. 1–52). For example, on the basis of his research on imitation in infants and young children over the past 20 years, Meltzoff (2005, p. 55) concluded that, ‘‘Imitation evolved through Darwinian means but achieves Lamarckian ends. It provides a mechanism for the ‘inheritance' of acquired characteristics. Imitation is powerful and can lead to rapid learning; it is essentially no-trial learning.''

Presently, research on imitation in infants and young children is multidisciplinary in its approach: Some developmental psychologists have suggested that imitation is goal-directed (Williamson & Markman, 2006; Wohlschlager, Gattis, & Bekkering, 2003) whereas the neuropsychologists Fadiga, Fogassi, Pavesi, and Rizzolatti (1995) have argued that imitation is mediated by a ‘‘mirror neuron'' system—a supramodal neural network (which may be functional from birth) linking perception and action in humans. In a recent review of the developmental and comparative literatures on imitation, Want and Harris (2002) have argued that imitation must be distinguished, procedurally and conceptually, from the ‘‘lesser'' forms of social learning that may lead to behavioral matching (see also the peer commentaries on that paper).

By the time Meltzoff had broken with Piagetian lore to provide evidence that human infants imitate from birth (Meltzoff, 1990; Meltzoff & Moore, 1977; 1983; but see Anisfeld, 1996; 2005), behavior analysts who took the opposing view that imitation is learned had already conducted a wealth of studies investigating the determinants of this repertoire. Unfortunately, and despite this early lead in investigating how imitation is acquired and maintained in children, studies drawn from

We are grateful to the children, their parents, and the staff at the Daycare Nursery and Center for Child Development at Bangor University for their participation in the study. The research reported here was supported by a PhD studentship awarded by the U.K. Economic and Social Research Council to Mihela Erjavec.

Requests for reprints should be addressed to Pauline J. Horne, School of Psychology, Bangor University, Brigantia Building, Penrallt Road, Bangor, Gwynedd, LL57 2AS United Kingdom (e-mail: p.j.horne@bangor.ac.uk).

doi: 10.1901/jeab.2008.89-183

this body of research are rarely (if ever) cited in the cognitive developmental literature. These behavior analytic research findings, as reviewed by Baer and Deguchi (1985), and more recent studies in infants (Poulson, Kymissis, Reeve, Andreatos, & Reeve, 1991; Poulson, Kyparissos, Andreatos, Kymissis, & Parnes, 2002), suggest that children show rapid and untrained acquisition of matching relations, termed *generalized imitation*, from an early age. In the procedure employed to test for generalized imitation, children are first trained to match a number of baseline behaviors that are modeled by the experimenter, initially under continuous reinforcement, and then under intermittent reinforcement. Next, models of target behaviors are presented on trials interspersed with the baseline models, matching responses to which continue to be intermittently reinforced; however, no consequences are delivered for responses to target models. Infants and older children, including those with developmental disabilities, are reported to match the target behaviors that are presented during generalized imitation tests. However, some researchers have noted that the topography of the trained baseline relations appears to determine whether or not the children match the unreinforced targets (Baer, Peterson, & Sherman, 1967; Bandura & Barab, 1971; Garcia, Baer, & Firestone, 1971; Poulson et al., 2002; Sherman, Clark, & Kelly, 1977; Steinman, 1970; Young, Krantz, McClannahan, & Poulson, 1994). For example, if the trained topographies were arm movements, target models of leg movements were less likely to be matched (Steinman, 1970). This led Baer and Deguchi (1985, pp. 183–184) to speculate that there may be subclasses of imitation rather than one overarching response class, a finding that has received support in studies of imitation in normally developing infants (Poulson et al., 2002) and children with developmental delay (Bandura & Barab, 1971; Garcia et al., 1971; Young et al., 1994). Nevertheless, within a topographical subclass, the evidence suggests that generalized imitation occurs reliably, even in infants.

One explanation for such nonreinforced target matching, put forward by Baer and colleagues (Baer & Deguchi, 1985; Baer et al., 1967), is the *conditioned reinforcer hypothesis*, which proposes that, in the course of learning matching relations, children also gradually learn to discriminate the topographical similarity (or parity) between the modeled stimulus and any of their responses that resemble it. This property of similarity acquires conditioned reinforcing properties that help to establish and maintain new matching relations that are not explicitly reinforced (Baer & Deguchi, 1985, pp. 199–217). There is indeed evidence to show that, from early infancy, parents begin to train matching responses during interactions termed *modeling frames* (Kaye, 1982; Pawlby, 1977) in which they imitate their infants' behaviors and establish sequences in which parent and child alternately produce the same behavior, and in the course of other matching games (Kokkinaki, 2003; Kokkinaki & Vasdekis, 2003). The key question is at what point in children's progress with trained matching does this conditioned reinforcement mechanism emerge? The studies in human infants conducted by Poulson et al. (1991, 2002) suggest that infants show generalized imitation towards the end of their first year. However, more recently, Horne and Erjavec (2007) have argued that, because the latter studies did not establish at the outset whether the target behaviors were novel (did not already feature in the infants' trained matching repertoires), they may not provide reliable evidence of novel matching relations acquired without direct reinforcement. When only target behaviors that did not feature in the infants' trained repertoires were presented in imitation tests, Horne and Erjavec found no evidence of generalized imitation, even after motor skills training of these novel target behaviors in a nonmatching context, followed by further imitation test trials. They concluded that infants well into their second year showed no reliable evidence of generalized imitation of novel target behaviors and that the infants' performances in the tests were more consistent with a trained matching account, as put forward by Skinner (1953, pp. 119–120): The infants' responses on test trials consisted mostly of well-practiced behaviors that commonly feature in infants' trained matching repertoires (such as tapping the backs of their hands and hugging gestures); these incorrect responses were emitted indiscriminately on most trials and bore only minimal resemblance to the modeled target gestures (such as touching a shoulder, elbow, or foot). Skinner's account predicts that similar errors—only partial matching responses to models of novel target behaviors, via generalization from the trained matching repertoire—should also be observable in older children, at least until their repertoires of trained matching relations become very fine grained and (or) parity-driven matching (as envisaged by Baer & Deguchi, 1985; and see Horne & Erjavec, 2007, pp. 64–65) emerges.

The present study aims to track the development of children's matching skills beyond infancy by investigating the performance of 2-year-olds and 3-year-olds in the standard generalized imitation test paradigm employed in previous behavior analytic studies. Clearly, it is not possible to measure the full extent of each young child's matching repertoire. Accordingly, the target gestures selected for investigation were variations on arm movements that were found to give rise to characteristic matching errors in young children (Bekkering & Wohlschlager, 2002; Bekkering, Wohlschlager, & Gattis, 2000; Gleissner, Meltzoff, & Bekkering, 2000). For example, when the 3- to 6-year-old children who participated in those studies were presented with hand-to-ear or hand-to-knee models, they often touched their ear(s) or knee(s), respectively, but did not always match the movement that the modeler used to do this. The most frequent errors occurred when the child was shown a target body-part touched by the modeler's *contralateral* hand but responded with an ipsilateral hand movement. These incorrect responses occurred on 60% of such trials (Gleissner et al., 2000) as compared with 10% of trials on which an *ipsilateral* hand-touch to target body-part gesture was modeled. The authors interpret these robust contralateral errors (Benton, 1959; Gordon, 1923; Head, 1920, 1926; Kephart, 1971; Wapner & Cirillo, 1968; also reported in 8-year-old children by Schofield, 1976) as evidence that children's matching performances are predominantly driven by their inferences about the primary goals of a modeled behavior; children focus on reproducing the primary goal—touching the correct body part—and attend less to the subsidiary goal(s) of how the touch is to be achieved. In support of goal theory, the authors have shown that when the children were required to match a model performing a contralateral reach near to, but not touching, the target body part, matching errors were reduced by 10%. These matching errors are particularly interesting to us because touches to some body parts are trained early in the course of social interaction and, during this training, modeling of ipsilateral touches is the norm. Children's errors occur mostly for contralateral behaviors (e.g., cross body arm movement to ear) that differ little in terms of topography from the ipsilateral counterparts. The data from these cognitive developmental studies are also intriguing because infants can be readily trained to match cross body gestures (Erjavec, 2002; Horne & Erjavec, 2007; Poulson et al., 2002) showing that cross body gestures per se are not difficult for infants and young children to perform.

However, these cognitive studies did not employ the test paradigm that is routinely employed in behavior analytic studies—there were no baseline matches trained in the experimental context and all children's responses to the target behaviors, whether correct or incorrect, were followed by verbal praise and other social reinforcement. In order to determine whether similar patterns of errors would occur in a generalized imitation test paradigm, some of the target gestures we presented incorporated the same basic arm movements as in the studies conducted by Bekkering and colleagues, but here they were interspersed with intermittently reinforced baseline matches. Given that a Skinnerian account predicts that children's histories of trained matching may play a role in the pattern of errors observed in these studies, we also included some target gestures that were identified by the first author (by observation and in consultation with the nursery nurses) as being commonly trained in the course of day-to-day matching games and other play procedures in the nursery that our participants attended, whereas other targets, matched for difficulty and body part touched, were defined a priori as uncommon target gestures because they were neither established during nursery games nor observed in the course of the children's free play. The first aim of the study was to replicate, using a generalized imitation test paradigm, the effects reported by Bekkering et al. (2000) and Gleissner et al. (2000) with respect to older children's responses to models of ipsilateral and contralateral touches to body parts. The second aim was to determine whether children's trained-matching repertoires contribute to these kinds of errors. The third aim was to determine whether cognitive goal theory or a trained matching account better explains children's responses to a set of gestures that were designed to tease apart these two theoretical positions.

Because our earlier research suggests that on generalized imitation test trials the responses of infants in their first and second year of life consisted entirely of gestures that featured in their previously directly trained matching repertoires, we expected to find that 2- to 3-year-old children's matching of commonly trained models would be superior to their matching of uncommon models. If children's matching responses are strongly determined by generalization of trained matching, we also expected that commonly trained gestures would be emitted in response to uncommon models much more often than vice versa. Second, we assessed whether children's responses to target models showed evidence of parity-based matching; this was done through the analysis of all incorrect responses and of changes in children's responding over repeated modeling trials. If children's matching were parity-driven, we would expect to observe increases in target matches (or better approximations to target responses) over trials in which corresponding models were repeatedly administered, especially across all uncommon target gestures.

Twenty children, 10 of whom were between 25 and 29 months old (younger group) and 10 between 35 and 41 months old (older group) at the start, participated in an AB procedure. In the baseline phase, four extraexperimentally trained matching relations were overtrained, first under continuous and then under intermittent reinforcement. In the imitation test phase of the experiment, over 12 testing sessions, the experimenter presented repeated modeling of target gestures, interspersed with intermittently reinforced baseline behaviors. The present design combined single-subject methodology, similar to that used in other behavior analytic studies of generalised imitation in young children, with an extensive replication over 20 subjects, commonly associated with the group designs employed by cognitive developmental researchers (e.g., Bekkering et al., 2000). This mixed design provided more extensive data than would result from either of these procedures alone.

METHOD

Subjects

Two groups of 10 children participated, with equal numbers of boys and girls. In the younger group (2-year-olds), the mean age was 27 months at the start of experiment and 28 months at the end; in the older group (3-year-olds), the mean age was 37 months at the start and 38 months at the end. The children attended the Daycare Nursery and Center for Child Development at Bangor University at least 2 days per week, and were recruited to the study by parental consent. The nursery staff reported that the children were developing normally; this is consistent with their performance on the Griffiths Mental Development Scales (GMDS; Griffiths, 1954), which was assessed for 12 out of 20 subjects. Table 1 shows the children's ages at the start and at the end of the study. All subjects completed the main procedure.

Setting and Apparatus

Experimental sessions were conducted in a research room in the daycare center. Two wall-mounted digital video cameras were employed; one recorded the behavior of the child and the other of the experimenter. A hidden radio microphone recorded the sounds. The outputs from the two cameras and the microphone were jointly incorporated in split-screen video recordings. Situated in a separate audio-visual suite, JVC SR-VS10 VHS/DV recorders provided the slow-motion and frame-by-frame viewing facilities required to code the target behaviors.

The experimenter sat on a beanbag, facing and at the same eye level as the child who was seated on a child-sized chair; a low table was placed between the child and experimenter. Age-appropriate toys, books, and stickers, used for warm-up play and as reinforcers, were concealed in a lidded toy chest until one of them was scheduled for presentation.

Stimuli

The visual stimuli employed were manual gestures performed by the experimenter.

Table 1

Subjects' group membership, genders, and ages at the start and end of the experiment.

Group	Child	Gender	Age (months/days)	
			At start	At end
Younger:	Cori	F	25/04	25/22
2-year-olds	Hari	M	25/09	26/05
	Jam	M	26/05	27/03
	Fin	M	26/06	29/04
	Ette	F	26/15	28/12
	Isa	F	27/16	30/05
	Non	F	28/11	30/13
	Jack	M	28/27	29/22
	Cal	M	29/02	29/20
	Mol	F	29/02	29/24
Older:	Cai	M	35/25	39/09
3-year-olds	Sion	M	36/05	37/04
	Cat	F	36/08	36/25
	Ann	F	36/11	37/00
	Fion	F	36/12	37/03
	Ady	M	36/13	37/00
	Stef	M	36/15	37/01
	Tom	M	36/25	39/29
	Ren	F	38/03	41/09
	Cara	F	41/29	42/14

Figure 1 (top panel) shows the four gestures used in baseline matching training (see Procedure). The remaining four panels in Figure 1 show the four sets of target gestures (20 gestures in all) that were employed in the imitation tests. The movement sequence that constituted each modeled baseline and target gesture is described in Table 2.

Target gestures: Set 1 and Set 2. The target gestures for Set 1 were touches to ear(s), and for Set 2, to shoulder(s), that were performed either uni- or bimanually, and with either ipsi- or contralateral hand movement. Within these two target sets, bimanual ipsilateral touches to ears and to shoulders (T5 and T11) were also identified as ''common'' targets—likely to be already trained matching responses for the subjects—because they feature in a singing-and-matching game of ''Heads and shoulders'' that was played routinely in the daycare center; this was also the case for one of the baseline gestures, B3 (both hands placed on head). The remaining bimanual target behaviors from Sets 1 and 2 (T6 and T12) are contralateral gestures, which are not presented in any nursery matching game, and were designated as ''uncommon'' target models. During a second kind of matching game, carers named and touched a number of different parts of their bodies (e.g., ear(s), nose, head, shoulders) and asked the children to do likewise, delivering praise when the children succeeded in doing so. Given that the carers usually modeled ipsilateral body touches in the course of the game, we also considered it likely that our subjects' matching repertoires would include unimanual ipsilateral touches to ear and shoulder (T1, T3, T7, and T9); these gestures were therefore also identified as putative common targets. The contralateral unimanual gestures (T2, T4, T8, and T10) were designated as uncommon targets because they did not feature in any of the nursery games.

Target gestures: Set 3 and Set 4. The target gestures for Set 3 were touches to mid-arm, and Set 4 to lower-arm locations, all performed unimanually and contralaterally, and incorporating movement that terminated at either top- or underside-of-arm locations. Within Set 3, underside-of-arm responses (T14 and T16) were identified as common targets, because they too feature in the naming-and-matching games described earlier. The remaining gestures from Set 3 (T13 and T15) were designated as uncommon targets, because they did not feature in the nursery games. Conversely, in Target Set 4, the top-of-arm responses (T17 and T19) were identified as common targets, because they feature in matching games such as ''pat-a-cake'' and mock punishment routines; these gestures were observed frequently, in our earlier research, in 1-year-old infants (see Introduction, and Horne & Erjavec, 2007). The remaining gestures from Set 4 (T18 and T20), which do not feature in these games, were designated as uncommon targets.

Procedure

Familiarization. The experimenter first established a good rapport with the children during unstructured daily play sessions in the nursery playroom. Familiarization play was age-appropriate, with an emphasis on turn-taking, but did not include any matching games. After several weeks the child was invited to participate in one-to-one play with toys in the test room. In this and subsequent experimental phases, children were tested each day they attended the nursery, with inevitable breaks for illnesses and holidays; each session lasted approximately 15 min, and most children completed the procedure in about one month. In all phases of the

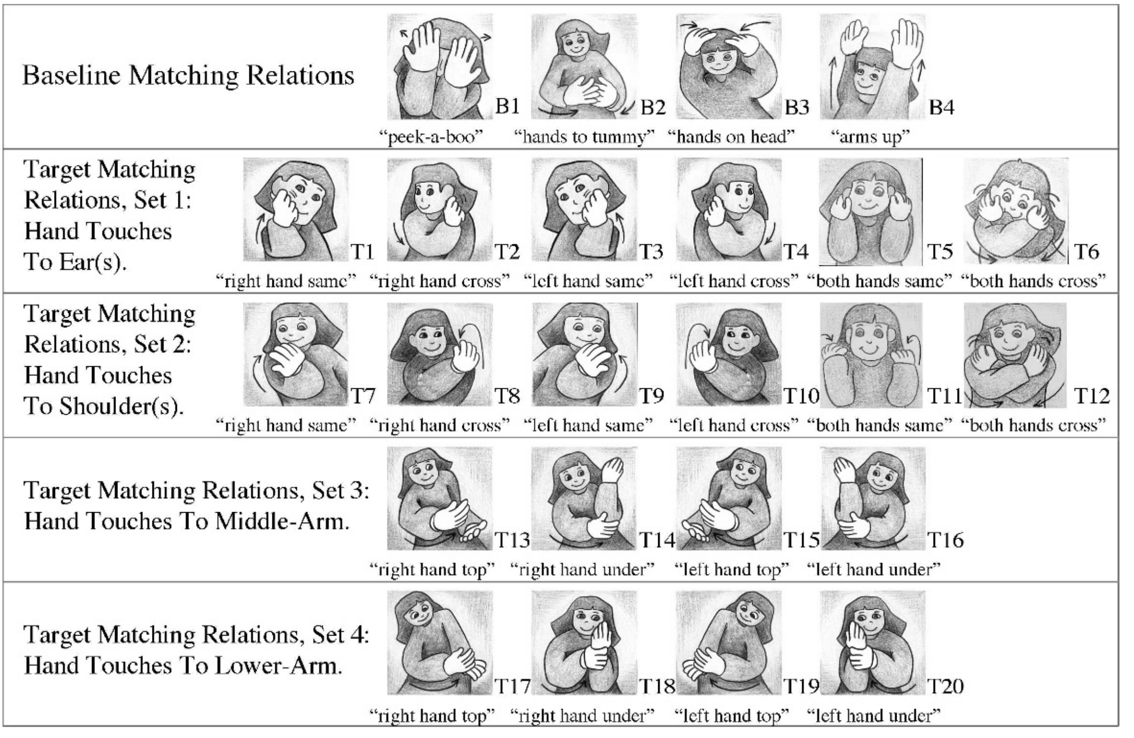

Fig. 1. Four baseline gestures (B1-B4) that featured in the subjects' trained baseline matching relations, and the 20 target gestures (T1-T20) employed in the imitation test.

experiment, a child was to be returned to the playroom if he or she showed discomfort at any point in the procedures. This seldom happened; all interrupted sessions were completed at the next testing opportunity. Whenever possible, if a child was happy and responsive, then more than one session was run in a single day (e.g., the first in the morning, and the second in the afternoon). All sessions ended with play.

Baseline matching training. For each child, matching responses to the four baseline gestures were trained to criterion. Following warm-up play, as in the previous condition, the experimenter asked the child, ''Shall we play our game?'' At the start of each modeling trial, the experimenter established eye contact with the child, for example by saying, ''[Name], look at me!'' Next, the experimenter asked, ''Can you do this?'' as she modeled one of the baseline gestures, then looked at the child expectantly. If the child produced a matching response the experimenter clapped enthusiastically exclaiming, ''Yeah!'' or ''Well done!'' and immediately delivered a sticker or a toy. If the child did not respond within 3 s, the experimenter again prompted him or her to do so by saying, ''You do it!'' or ''Show me!'' If the child still produced no response, the model and prompt sequence was repeated for the same gesture, up to two more times. If no response was emitted over three such prompts, the experimenter shaped the target response by gently moving the child's hands into a matching response configuration (''putting through''—see Baer & Deguchi, 1985, p. 182) then delivered social praise followed by a sticker or a toy. Nonmatching responses were corrected in a similar manner; the experimenter said, ''Not quite; this is how we do it!'' as she manually guided the child's correct response, then delivered the reinforcers. In each training session there were three trials of each of the four baseline gestures (12 trials per session), with up to three models per trial (as necessary). The modeled gestures were presented in a predetermined randomized order, with the added constraint that no more than two trials of the same gesture could occur in succession. The criterion was five out of six,

Table 2

For each baseline (B) and target (T) gesture, description of movements modeled by experimenter and response variations that met the matching criteria.

Baseline / target gestures	Behavior modeled by experimenter	Response variations accepted as correct
B1 Peek-a-boo	Both hands covering eyes, opening with a ''boo'' sound	Both hands on any part of face
B2 Hands to tummy	Both hands tapping tummy	Both hands touching tummy/chest
B3 Hands on head	Both hands placed on top of head	Both hands on head, touching hair
B4 Arms up	Both arms raised above head, stretching and looking up at hands	Both hands up, at head level or above
*T1 Right hand same ear	Right hand pulling right ear (ipsilateral movement)	Right hand touching right ear
T2 Right hand cross ear	Right hand pulling left ear (contralateral movement)	Right hand touching left ear
*T3 Left hand same ear	Left hand pulling left ear (ipsilateral movement)	Left hand touching left ear
T4 Left hand cross ear	Left hand pulling right ear (contralateral movement)	Left hand touching right ear
*T5 Both hands same ear	Both hands simultaneously pulling both ears (ipsilateral movement)	Both hands touching both ears, without crossing body midline
T6 Both hands cross ear	Both hands simultaneously pulling both ears (contralateral movement)	Both hands touching both ears, crossing body midline
*T7 Right hand same shoulder	Right hand touching top of right shoulder (ipsilateral movement)	Right hand touching top, side, or front of right shoulder
T8 Right hand cross shoulder	Right hand touching top of left shoulder (contralateral movement)	Right hand touching top, side, or front of left shoulder
*T9 Left hand same shoulder	Left hand touching top of left shoulder (ipsilateral movement)	Left hand touching top, side, or front of left shoulder
T10 Left hand cross shoulder	Left hand touching top of right shoulder (contralateral movement)	Left hand touching top, side, or front of right shoulder
*T11 Both hands same shoulder	Both hands simultaneously touching tops of both shoulders (ipsilateral movement)	Both hands touching tops, sides, or fronts of both shoulders, without crossing body midline
T12 Both hands cross shoulder	Both hands simultaneously touching tops of both shoulders (contralateral movement)	Both hands touching tops, sides, or fronts of both shoulders, thus crossing body midline
T13 Right hand top mid-arm	Right hand touching top of resting left mid-arm (contralateral)	Right hand touching top of resting left arm around middle
*T14 Right hand under mid-arm	Right hand touching underside of raised left mid-arm (contralateral)	Right hand touching underside of raised left arm around middle
*T15 Left hand top mid-arm	Left hand touching top of resting right mid-arm (contralateral)	Left hand touching top of resting right arm around middle
T16 Left hand under mid-arm	Left hand touching underside of raised right mid-arm (contralateral)	Left hand touching underside of raised right arm around middle
*T17 Right hand top low-arm	Right hand touching top of resting left low-arm (contralateral)	Right hand touching top of resting left arm around lower arm
T18 Right hand under low-arm	Right hand touching underside of raised left low-arm (contralateral)	Right hand touching underside of raised left arm around lower arm
*T19 Left hand top low-arm	Left hand touching top of resting right low-arm (contralateral)	Left hand touching top of resting right arm around lower arm
T20 Left hand under low-arm	Left hand touching underside of raised right low-arm (contralateral)	Left hand touching underside of raised right arm around lower arm

Note: Targets identified as *common* are marked with asterisks.

unguided, correct responses to each of the four baseline models, over two consecutive sessions. When performance met the 100% reinforcement criterion, the reinforcement rate was reduced to 50% on a VR2 schedule. The intermittent reinforcement criterion was 11 out of 12 correct responses across three trials per gesture within a single session.

Imitation test. For each child, the gestures in the four target sets (see Figure 1) were divided quasirandomly and without replacement into four test groups, each of which contained a unimanual touch to ear, a unimanual touch to shoulder, a bimanual touch to either shoulder or ear, a touch to middle arm, and a touch to lower arm (see Figure 1). The five target gestures in each test group were presented in each of three consecutive sessions; each child therefore completed 12 test sessions in all. In each test session, the five target gestures were interspersed with the four baseline gestures, in a prerandomized sequence. Each gesture (baseline and target) was presented twice, making 18 trials per session in total. Correct responses to the baseline gestures continued to be followed by intermittent delivery of reinforcers, but there were no scheduled consequences (reinforcement or correction) for responses to target models. The criterion for performance of the baseline gestures was 14 out of 16 (88%) correct over two consecutive sessions. If the criterion was not met, baseline responding was to be re-established before the next testing session was conducted; however, this was not necessary for any of the children. All gestures were presented in the manner described earlier: The experimenter prompted the child to ''Do this'' and presented the modeling of each gesture up to three times per trial, as necessary. The experimenter smiled throughout the modeling and response periods of each trial in order to ensure that responding in target gesture trials was not suppressed, as may have been the case if such trials were selectively followed by a ''still face'' expression by the experimenter (Striano & Lizkowski, 2005).

Developmental test. After the imitation tests were completed, an accredited administrator of the Griffiths' Mental Development Scales assessed the children's development in terms of the GMDS General Quotient; this developmental test was conducted for the 12 children who continued to attend the Nursery and whose parents had consented to this part of the procedure. At the end of the procedure, the children's parents were given videotapes containing representative recordings of their children's test sessions, a letter that fully explained the procedures and results of the study, and the Griffiths' test report for their child; each child chose and received a soft toy and a sticker book.

Coding

The modeled baseline and target behaviors, and the corresponding response criteria that were formulated prior to the start of the study, are shown in Table 2. The response criteria excluded behaviors commonly produced by children of this age, such as pointing, rubbing eyes, mouthing fingers, clapping, and speaking to the experimenter.

Children's responses to baseline models were classified as either *baseline matches, incorrect,* or *no response.* Likewise, children's responses to target models were broadly classified as either *target matches, incorrect,* or *no response.* Across unimanual models (16 of the 20 target gestures), all target matches were further classified as either *correct,* if a child responded with the same hand as the modeler to produce the correct response configuration, or *mirror,* if a child responded with the opposite hand and arm to produce a mirror-image of the modeled gesture (naturally, this classification was unnecessary for target matches performed with both hands). All other responses were coded as incorrect and further classified as *target mismatches, baseline mismatches,* or *other behaviors.* The form of each such gesture, the number of models (1, 2, or 3) per trial required to evoke a response, and whether reinforcement was given in a particular baseline gesture matching trial, were also recorded.

Two undergraduate researchers, with experience in working with preschool children but naïve to the experimental hypotheses, scored a randomly selected 30% of the data. Inter-observer agreement was calculated on a point-to-point basis, by dividing the number of agreements by the total number of agreements and disagreements, and multiplying the result by 100. Inter-observer agreement was 98% over 828 baseline trials and 96% over 720 target trials.

RESULTS

Baseline Gestures

Baseline matching training. The 2-year-old children completed baseline training under continuous reinforcement within three to five sessions, and the 3-year-olds in three to six sessions. The children's performances met the intermittent reinforcement criterion in three to six sessions; there was no evidence that the 3-year-olds learned faster than the 2-year-olds.

Baseline matching during imitation tests. Mean correct baseline matching for the younger group was 93% (range: 92–99%) and for the older group 96% (range: 93–100%). Baseline gestures had to be modeled more than once on a mean of 2% of trials (range: 0–6%) for the younger group and 4% of trials (range: 0–9%) for the older group.

Target Gestures: General Properties of Responses to the 20 Target Models

Prompts. No responses were recorded on only four target trials, two each for a 2-year-old and a 3-year-old). More than one model was required to evoke a response on only 2% of target trials for the 2-year-olds (range: 0–12%) and on 4% of target trials for the 3-year-olds (range: 0–11%). The procedure was therefore very effective in evoking these young children's responses to target models presented over the 12 test sessions.

Matching responses to target models: age differences. In order to calculate the frequency with which the children produced the same response configuration as the target model, the lateral difference between correct and mirror responses was ignored and mirror responses were included in the matching response category. Using this measure, the children produced matching responses on just over half of the target trials: The 2-year-olds matched on 44% of all target trials (range: 31–61%), and the 3-year-olds on 60% of trials (range: 47–74%); statistically, this difference across age groups was significant (Mann-Whitney $U = 17.5$, $p = .01$). Separate analyses for each stimulus set show that the 2-year-olds matched on 41% of Set 1 trials (range: 17–58%), on 43% of Set 2 trials (range: 28–56%), on 45% of Set 3 trials (range: 0–79%), and on 49% of Set 4 trials (range: 8–83%); statistically, their matching performances did not differ between target gesture sets (Friedman's $\chi2 = 2.24$, $df = 3$, $p = .53$). The 3-year-olds matched on 53% of Set 1 trials (range: 33–81%), on 60% of Set 2 trials (range: 36–78%), on 76% of Set 3 trials (range: 50–96%), and on 51% of Set 4 trials (range: 8–88%); statistically, their matching performances differed between target gesture sets (Friedman's $\chi2 = 12.91$, $d = 3$, $p = .01$). These data show that even when mirror responses are counted as correct, the children produced a high frequency of errors. Overall, the 3-year-old children matched the target models more frequently than the younger children. Whereas the younger children matched gestures from the four target sets approximately equally, the older children performed better on some sets than on others.

Correct responses versus mirror responses. Previous studies have reported that kindergarten-age children, seated face-to-face with an adult modeler, produce mirror matching responses more frequently than correct matching responses; even adults produce mirror responses in some such tasks (see Wohlschlager et al., 2003). Each child's correct and mirror responses to the 16 unimanual target behaviors modeled in the present study are shown in Figure 2a (2-year-olds) and Figure 2b (3-year-olds). Overall, the 20 children produced mirror responses on 41% of trials (range: 17–64%) and correct responses on only 10% of trials (range: 0–25%). All 20 children performed more mirror responses than correct responses (Wilcoxon's $T = 0$, $p = .0001$). Figure 3 shows that both age groups produced more mirror matches than correct matches for the unimanual gestures in all four target sets: All eight plotted effects—percentages of trials with correct versus mirror matches for each age group and each gesture set—are statistically significant (with p value range .0001–.03). This preponderance of mirror over correct responding is consistent with the developmental literature and extends those findings to the generalized imitation test paradigm.

Changes in accuracy of children's matching over trials. If the children's matching performances were parity-driven, we would expect the accuracy of their target matches to improve over repeated modeling trials. Therefore, the development of each child's matching responses to each target gesture was examined by comparing the number of correct responses on the first three modeling trials to those on the last three modeling trials. Thus, for each gesture, a child's total (0–3) of target matches on the first three trials was subtracted from his or her total (0–3) target matches on the last three trials. The resulting scores were then used to classify the child's performance for each target gesture as either improving (a positive score), getting worse (a negative score), or remaining unchanged (a zero score). These scores, showing changes in matching performance over trials for each target gesture and for each child, are shown in Table 3.

For the 2-year-olds as a group, matching improved across trials for 17% of target

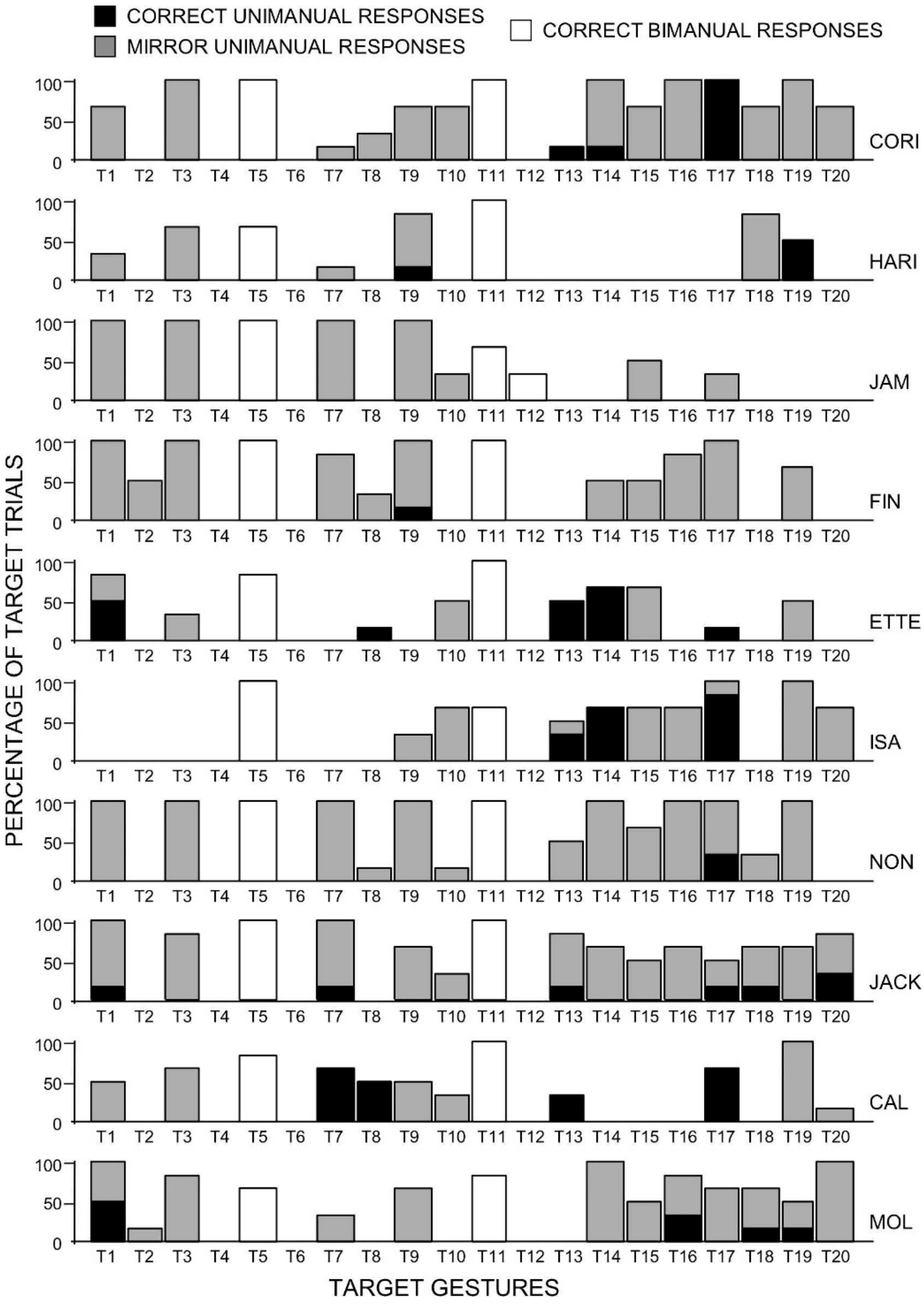

Fig. 2a. For each 2-year-old child, percentage of target trials on which correct and mirror responses were emitted to each of 20 target gestures (T1-T20).

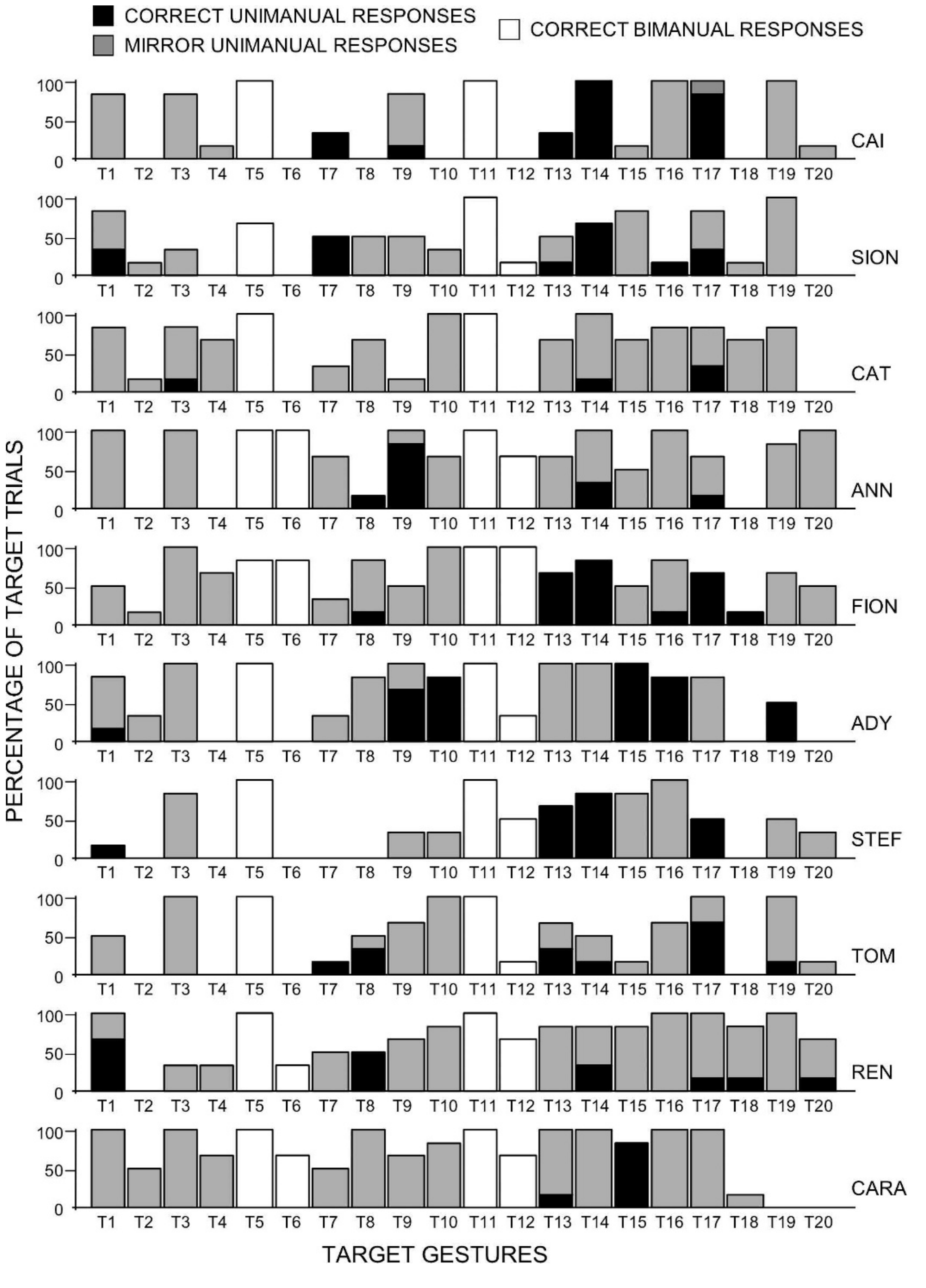

Fig. 2b. As in Fig. 2a, for each 3-year-old child.

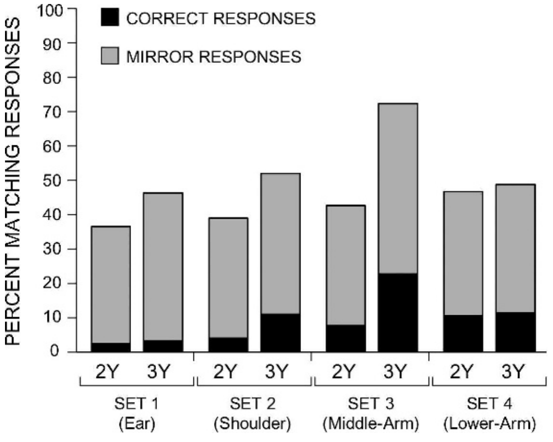

Fig. 3. Percentages of trials on which correct and mirror responses were produced by the 2-year-olds (2Y) and the 3-year-olds (3Y) to unimanual Set 1 (T1-T4) and Set 2 (T7-T10) target gestures, and Set 3 (T13-T16) and Set 4 (T17-T20) target gestures.

gestures (range: 5–30%), got worse for 15% of target gestures (range: 5–25%), and did not change for the remaining gestures; statistically, the children's matching was no more likely to improve than to get worse over trials (Wilcoxon's $T = 16$, $p = .43$). For the 3-year-olds, matching improved across trials for 19% of the target gestures (range: 5–40%), got worse for 24% (range: 15–45%), but there was no change for the remaining gestures (range: 30–70%); statistically, their matching performances were no more likely to improve than to get worse over trials (Wilcoxon's $T = 15.5$, $p = .40$). The data do not appear to support the hypothesis that parity serves as a conditioned reinforcer in these young children's matching performances (see Table 3).

Hand bias: right- versus left-handed responses. Although the unimanual target gestures were modeled with the right and left hand on an equal number of trials in each target set, right-handed responses were more frequent than left-handed responses across all unimanual trials. The 2-year-old children responded with right hand gestures on 52% of trials (range: 29–78%), with left hand gestures on 31% of trials (range: 4–53%), and bimanually on the remaining trials. Statistically, the difference between right- and left-handed responses approached significance (Wilcoxon's $T = 7$, $p = .07$).

The 3-year-old children responded with right-hand gestures on 57% of trials (range: 21–77%), with left-hand gestures on 31% of trials (range: 0–75%), and bimanually on the remaining trials. Statistically, the difference between right- and left-handed responses was significant (Wilcoxon's $T = 8$, $p = .05$). It appears that idiosyncratic biases in handedness also determine young children's responses to models of unimanual target behaviors.

Summary of Children's Matching Performances over the Four Target Sets

Following baseline matching training, all children continued to match baseline models at very high rates, under intermittent reinforcement, throughout the main testing phase of the experiment. On target behavior test trials, when mirror responses are counted as correct, the children matched on just over half of these trials (see Figure 3). The accuracy of children's matching did not improve over repeated trials. Younger children produced fewer target matches than the older children, but the overall pattern of results was similar for the two age groups. When mirror responding is analyzed separately, the children produced more mirror responses than correct matching responses (see Figure 3). Most children emitted right-handed responses more frequently than left-handed responses.

Target Set 1 and Set 2: Touches to Ear(s) and to Shoulder(s)

Set 1 and Set 2 touches to ear(s) and shoulder(s), respectively, were modeled either uni- or bimanually, and incorporated either ipsi- or contralateral hand movement. The group frequencies of target and other responses produced by the 2-year-olds and the 3-year-olds to each target model are given in Figure 4. This shows that when the children were presented with Set 1 ear gesture models, they mostly touched one or both of their ears and only infrequently touched their shoulders or other body parts. Likewise, when they were presented with Set 2 shoulder gesture models, the children also mostly touched one or both of their shoulders, although incorrect responses that terminated on other body parts were comparably more frequent than for the Set 1 ear gestures. However, across both target sets, when the children touched the target location(s) they frequently did so with the wrong hand(s): They produced ipsilateral responses

Table 3

Changes in matching accuracy over trials for each named target gesture and for each child.

Group	Child	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20
Younger: 2-year-olds	Cori	0	0	0	0	0	0	−1	0	−1	+2	0	0	+1	0	0	0	0	+2	0	+2
	Hari	−2	0	0	0	+2	0	+1	0	−1	0	0	0	0	0	0	0	0	−1	0	0
	Jam	0	0	0	0	0	0	0	0	0	0	−2	+2	0	0	−1	0	−2	0	0	0
	Fin	0	−1	0	0	0	0	+1	+2	0	0	0	0	0	−1	−2	+1	0	−1	0	0
	Ette	−1	0	−2	0	+1	0	0	−1	0	+1	0	0	−3	0	+2	0	+1	0	+1	0
	Isa	0	0	0	0	0	0	0	0	0	+2	−3	0	−1	+2	+2	0	0	0	0	0
	Non	0	0	0	0	0	0	0	+1	0	+1	0	0	+1	0	−1	0	0	0	0	0
	Jack	0	0	−1	0	0	0	0	0	0	0	0	0	+1	0	−1	−2	−1	+1	+1	+2
	Cal	+3	0	+2	0	+1	0	−2	−3	+1	+2	0	0	0	0	0	0	0	0	0	+1
	Mol	0	−1	−1	0	0	0	−2	0	0	0	+1	0	0	0	−1	+1	0	+2	+1	−1
Older: 3-year-olds	Cai	−1	0	+1	−1	0	0	−2	0	−1	0	0	0	0	0	+1	0	0	0	0	−1
	Sion	+1	−1	0	0	0	0	+1	+1	+1	0	0	+1	+1	−2	+2	−1	−1	−1	+2	0
	Cat	+1	+1	−1	0	0	0	0	+1	+1	−1	0	0	0	0	0	0	−1	0	0	0
	Ann	0	0	0	0	0	0	−1	−1	0	0	0	+2	+1	0	0	0	−1	+1	+1	0
	Fion	+1	+1	0	+2	+1	0	0	−1	−1	0	0	+2	−2	−1	−1	+1	0	+1	−2	+1
	Ady	−1	0	0	0	0	0	+2	+1	0	−1	0	0	0	0	0	−1	+2	0	0	0
	Stef	−2	0	−1	0	0	0	0	0	0	0	0	0	+2	−1	+1	0	−2	0	+1	−1
	Tom	−1	0	0	0	0	0	−1	−1	0	0	0	0	−2	0	−1	+2	0	0	0	+1
	Ren	0	0	0	−2	0	−2	0	−1	0	−1	0	−2	−1	−1	−1	0	0	−1	0	+1
	Cara	0	−1	0	+2	0	0	+1	0	−2	+1	0	0	0	0	−1	0	0	−1	0	0

Note: Positive scores (+1 to +3) show improvement in matching over trials; negative scores (−1 to −3) show that matching got worse; zero scores show no change.

SET 1 (EAR)	CHILDREN'S RESPONSES TO TARGET MODELS (YOUNGER + OLDER)								
MODELED GESTURES	T1	T2	T3	T4	T5	T6	B1-B4	T7-T20	OTHER
T1	**7+8**	0+5	**37+37**	0+0	16+8	0+1	0+0	0+0	0+1
T2	35+41	**0+0**	1+3	**4+8**	19+7	0+0	0+0	0+1	1+0
T3	**44+48**	0+1	**0+1**	0+0	15+10	0+0	0+0	0+0	1+0
T4	16+13	**0+15**	34+19	**0+0**	9+12	0+0	0+0	0+0	1+0
T5	1+1	1+1	0+0	0+0	**54+57**	0+0	4+0	0+1	0+0
T6	4+0	0+0	0+0	0+0	43+38	**0+17**	0+0	0+0	13+5
TOTAL	*218*	*23*	*132*	*12*	*288*	*18*	*4*	*2*	*22*

SET 2 (SHOULDER)	CHILDREN'S RESPONSES TO TARGET MODELS (YOUNGER + OLDER)								
MODELED GESTURES	T7	T8	T9	T10	T11	T12	B1-B4	T1-T6 & T13-T20	OTHER
T7	**5+6**	7+12	**26+16**	0+3	21+16	0+0	1+0	0+3	0+3
T8	27+15	**4+7**	7+0	**5+23**	6+7	0+0	2+1	5+5	2+2
T9	**38+28**	1+1	**2+10**	1+2	11+17	0+0	2+0	2+0	2+2
T10	16+0	**18+36**	17+5	**0+5**	3+7	0+0	0+0	4+4	2+3
T11	0+0	0+0	0+0	0+0	**55+60**	0+0	4+0	0+0	2+0
T12	0+0	0+0	0+0	0+0	32+19	**2+25**	1+0	1+0	27+15
TOTAL	*135*	*87*	*83*	*39*	*253*	*27*	*11*	*24*	*60*

Fig. 4. The frequencies of children's responses to modeling of Set 1 ear touches (top panel) and Set 2 shoulder touches (bottom panel), for the two age groups. The cells with counts of target matches are shaded grey: Correct and mirror responses are presented in bold numerals; all other counts represent entirely incorrect responses. Common target gestures were T1, T3, and T5 in Set 1, and T7, T9, and T11 in Set 2; the remaining targets were designated as uncommon.

to contralateral models and bimanual responses to unimanual models on about half of all trials.

Target matches to ipsilateral (common) versus contralateral (uncommon) models. Figure 5 (top left panel) shows the percentages of trials on which the children produced target matches to the ipsilateral and contralateral models from Set 1 (ear) and Set 2 (shoulder), plotted separately for uni- and bimanual gestures, and for the two age groups. The children produced more matching responses to ipsilateral models than to contralateral models across both age groups, for both unimanual and bimanual gestures (all four effects shown in Figure 5 are statistically significant, with p value range .0001–.02).

Further analyses for each stimulus set showed that, overall, the children matched the ipsilateral (ear touch) models from Set 1 on 84% of trials (range: 50–100%), and the ipsilateral (shoulder touch) models from Set 2 on 75% of trials (range: 50–100%); statistically, this difference was significant (Wilcoxon's $T = 28.5$, $p = .01$). However, this difference in matching across Set 1 and Set 2 was not due to a general preference for touching ear(s) rather than shoulder(s), because the children's responses to contralateral models from the two target sets showed the opposite pattern: Overall, target matches to contralateral Set 1 targets were emitted on 13% of trials (range: 0–63%), and to Set 2 contralateral targets on 32% of trials (0–96%); statistically,

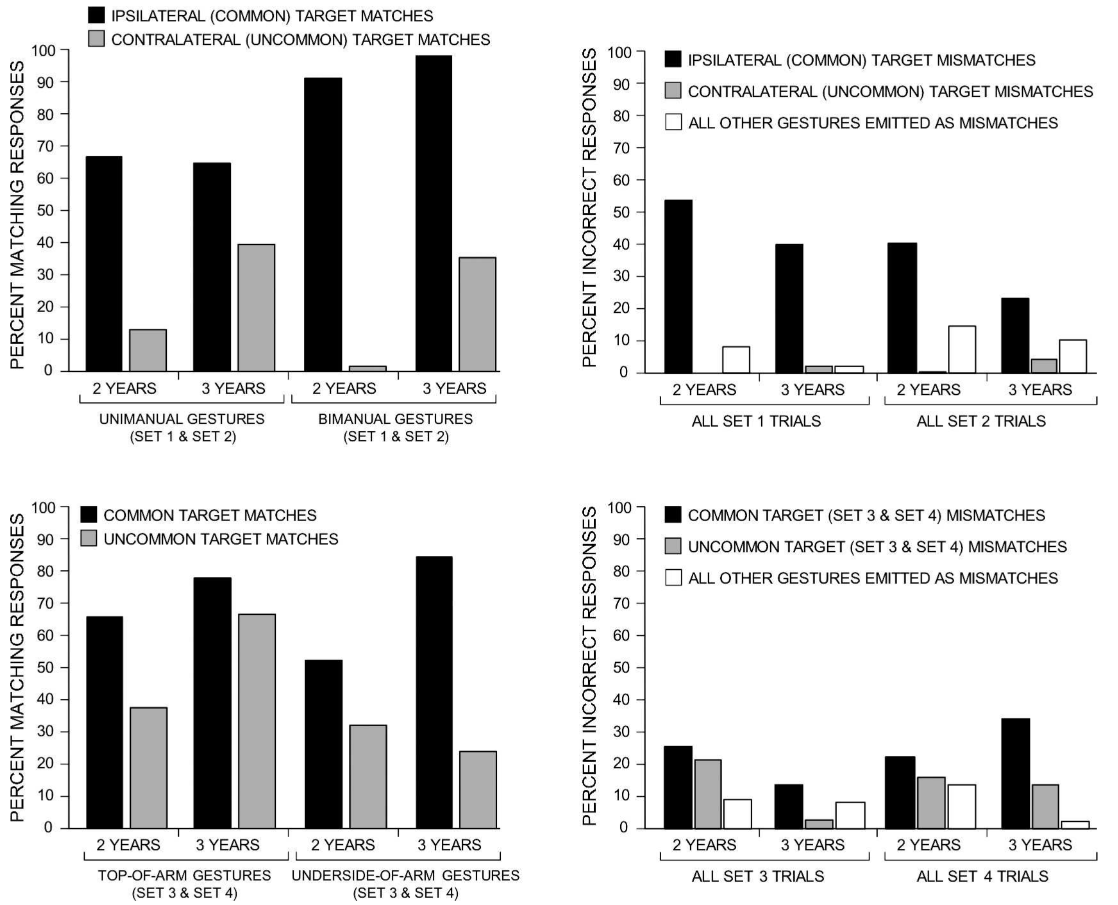

Fig. 5. Percentages of trials on which children in each age group produced target matches and target mismatches. For Set 1 and Set 2 gestures, target matches (upper left panel) are plotted for unimanual ipsilateral (common) targets, unimanual contralateral (uncommon) targets, bimanual ipsilateral (common) targets, and bimanual contralateral (uncommon) targets; upper right panel shows frequencies of ipsilateral (common) gestures, contralateral (uncommon) gestures, and all other responses, emitted as mismatches across all Set 1 and Set 2 trials. For Set 3 and Set 4 gestures, target matches (bottom left panel) are plotted for top-of-arm common targets, top-of-arm uncommon targets, underside-of-arm common targets, and underside-of-arm uncommon targets; target mismatches (bottom right panel) are plotted for common gestures, uncommon gestures, and all other responses.

this difference was also significant (Wilcoxon's $T = 9.5$, $p = .0001$).

Age differences. Figure 5 shows that the two groups of children matched ipsilateral (common) Set 1 and 2 models, on 74% of trials for the younger group (range: 33–100%) and 75% of trials for the older group (range: 56–94%); statistically, these scores were not different (Mann-Whitney $U = 47.5$, $p = .85$). However, there were age differences in children's responding to contralateral (uncommon) targets: The 2-year-olds matched on only 9% of contralateral trials (range: 0–17%) but the 3-year-olds did so on 38% of trials (range: 3–75%); this difference was statistically significant (Mann-Whitney $U = 10.5$, $p = .0001$).

Target matches to unimanual versus bimanual models. Figure 5 shows that, across Set 1 and Set 2 ipsilateral (common) target gesture trials, the children produced more target matches to bimanual than to unimanual models. Statistically, the children's matching of bimanual ipsilateral models (94% of trials) was significantly better than for unimanual ipsilateral models (65% of trials; Wilcoxon's $T = 5$, $p = .0001$).

Analysis of the children's responses to contralateral (uncommon) target gestures in both sets reveals an effect opposite to that just described for ipsilateral (common) target gestures and shows that there is no evidence of a general preference for bimanual responding: There were more target matches to unimanual (26% of trials) than to bimanual (18% of trials) contralateral (uncommon) models (Wilcoxon's $T = 52$, $p = .09$). Thus children in both groups performed best—most of them errorlessly—in response to bimanual ipsilateral (common) models; they also emitted matching responses on a substantial proportion of unimanual ipsilateral (common) trials. By contrast, they seldom emitted matching responses to unimanual contralateral (uncommon) models, and had the fewest correct responses—virtually none in the younger group—to bimanual contralateral (uncommon) models.

Children's matching errors: same-for-cross. Figure 5 (top right panel) shows, for each age group, the percentages of trials on which ipsilateral (common) target gestures, contralateral (uncommon) target gestures, and any other gestures occurred as incorrect responses to Set 1 and Set 2 models. This figure shows that, in each target set and in both age groups, the children produced ipsilateral (common) target gestures as mismatches much more frequently than contralateral (uncommon) targets and all other gestures: All eight comparisons shown in Figure 5 (of ipsilateral mismatches versus contralateral mismatches, ipsilateral mismatches versus all other errors, across both age groups) are statistically significant (with a p value range .0001–.03).

Children's matching errors: two hands for one hand. Further analyses revealed that children frequently produced bimanual ipsilateral (common) responses (T5 and T11) as mismatches to all unimanual target models (see Figure 4); these bimanual responses were recorded on 19% of unimanual Set 1 and Set 2 trials (range: 2–69%). However, unimanual target responses to bimanual target models occurred on only 2% of bimanual trials (range: 0–8%); this difference was statistically significant, with all 20 children emitting more bimanual-for-unimanual responses than vice versa (Wilcoxon's $T = 0$, $p = .0001$).

Children's matching errors: gestures from other target sets and other trained matching relations. Figure 5 also shows that gestures other than those from Set 1 and Set 2 were produced as mismatches on 9% trials (range: 0–19%); all younger children, and 8 out of 10 older children, produced some such responses (and see Figure 4). First, gestures in which hand(s) touched body parts other than ear(s) or shoulder(s) occurred on 4% of all Set 1 (ear) and Set 2 (shoulder) trials (range: 0–13%); these mismatches mostly consisted of well-established matching responses, such as clapping, touching eyes or head (B1 and B3), and tapping of back of hand (T17 and T19 from Set 4); 17 out of 20 children produced some such responses. Second, hug-like responses were produced solely to contralateral bimanual target models; they were recorded on 6% of T6 (hands cross to ears) trials (range: 0–34%) and on 27% of T12 (hands cross to shoulders) trials (range: 0–100%). These responses, which feature commonly in children's trained matching repertoires, were produced by 4 children (3 younger, 1 older) to Set 1 models, and by 12 children (6 younger and 6 older) to Set 2 models.

Summary: Set 1 and Set 2

The data for Set 1 and Set 2 show better matching of ipsilateral (common) than contralateral (uncommon) gestures irrespective of whether an ear touch or a shoulder touch was modeled (see Figure 5). Figure 5 also shows that ipsilateral (common) responses were emitted as mismatches on 54% of Set 1 trials and 40% of Set 2 trials by the younger children and on 40% of Set 1 trials and 24% of Set 2 trials for the older children. Whether the better matching observed for ipsilateral gestures is due to the relative difficulty of producing a cross-body movement as opposed to a same side movement cannot be determined from the Set 1 and Set 2 data because it was also noted that the ipsilateral gestures were all commonly trained in the nursery setting whereas the contralateral gestures were not. The role of prior training as opposed to movement across the body can be more readily adduced from the children's responses to the Set 3 and Set 4 models, all of which were contralateral but half of which were commonly trained matches.

Target Set 3 and Set 4: Touches to Middle- and Lower-Arm

Set 3 and Set 4 target touches to middle- and lower-arm locations, respectively, were per-

SETS 3 & 4 (MIDDLE- & LOWER-ARM)		CHILDREN'S RESPONSES TO TARGET MODELS (YOUNGER + OLDER)										
		T13	T14	T15	T16	T17	T18	T19	T20	B1-B4	T1-T12	OTHER
MODELED GESTURES	T13	**9+14**	1+0	**8+28**	11+1	6+10	1+0	11+3	7+0	0+0	3+4	2+0
	T14	5+1	**9+26**	6+0	**24+26**	6+0	2+0	5+1	0+0	2+0	1+3	0+3
	T15	**28+27**	4+0	**0+11**	1+0	13+14	3+0	1+1	5+0	2+1	2+5	1+1
	T16	9+4	**28+43**	6+1	**2+7**	0+1	6+1	0+1	0+0	3+1	1+0	5+1
	T17	0+5	0+0	1+0	4+0	**19+22**	2+0	**19+28**	1+0	2+0	4+1	8+1
	T18	4+4	5+0	5+0	7+6	8+12	**2+2**	6+24	**17+10**	0+2	0+0	5+0
	T19	8+11	0+0	0+4	0+0	**37+40**	5+0	**4+4**	0+0	0+0	0+1	6+0
	T20	11+5	2+7	0+1	0+1	12+23	**18+16**	7+5	**2+1**	3+0	1+0	3+1
TOTAL		*147*	*125*	*71*	*90*	*223*	*58*	*120*	*43*	*16*	*26*	*37*

Fig. 6. The frequencies of children's responses to modeling of Set 3 middle-arm touches (top half of table) and Set 4 lower-arm touches (bottom half of table), for the two age groups. The cells with counts of target matches are shaded grey: Correct and mirror responses are presented in bold numerals; all other counts represent entirely incorrect responses. Common target gestures were T14 and T16 from Set 3, and T19 and T20 from Set 4; all other targets were designated as uncommon.

formed unimanually and contralaterally, and incorporated movement that terminated at either top- or underside-of-arm locations; half of these gestures were identified as common targets (T14, T16, T17, and T19), and the remaining gestures were designated as uncommon targets (T13, T15, T18, and T20; see Stimuli, Figure 1, and Table 2).

Exact frequencies of different responses to modeling of each target gesture, produced by the children in each age group, are given in Figure 6. When the children were presented with Set 3 middle-arm gesture models, they mostly responded by touching middle-arm locations, but touches to lower-arm locations were also quite frequent, and most children also reached for their shoulders or other body parts. Likewise, the children most often responded to the Set 4 lower-arm gestures by touching lower-arm locations, but touches to middle-arm locations were also quite frequent, and most children also reached for other body parts. Overall, children were less likely to touch the correct body parts in Set 3 and Set 4 trials than they were in Set 1 and Set 2 trials. Across all Set 3 and Set 4 modeling trials, about half of all responses contained incorrect gesture configurations (wrong arm movements).

Target matches to common versus uncommon models. The percentages of trials on which the children produced target matches to common and uncommon Set 3 and Set 4 models, plotted separately for top- and underside-of-arm manual gestures, and for the two age groups, are presented in Figure 5 (bottom left panel). The children produced more matching responses to common models than to uncommon models over both types of gestures and in both age groups. Overall, target matches for top-of-arm gestures were recorded on 72% of common trials (range: 17–100%) and on 52% of uncommon trials (range: 0–100%); this difference was statistically significant (Wilcoxon's T = 46.5, p = .03). The differences in matching responses were even more prominent for underside-of-arm gestures, where target matches were recorded on 68% of common trials (range: 0–100%) and on only 28% of uncommon trials (range: 0–75%; Wilcoxon's T = 13, p = .0001. Further analyses showed that, overall, children's matching of uncommon target models was better for Set 3 than for Set 4 gestures (52% versus 28%; Wilcoxon's T = 40.5, p = .02). However, this effect was not due to a general preference for touching middle-arm rather than lower-arm locations, because the rates of target matches of common models from the two target sets were comparable (72% for Set 3 versus 68%

for Set 4; Wilcoxon's $T = 62.5$, $p = .78$). Overall, the children's matching of uncommon Set 3 and Set 4 target gestures was better than their matching of contralateral and uncommon gestures from Set 1 and Set 2 reported earlier (see Figure 5, top left panel).

Age differences. Figure 5 shows that the older children produced somewhat higher rates of target matches than the younger group, but these differences were not statistically significant either for common targets (82% for the older group, range: 67–100%, versus 59% for the younger group, range: 8–100%; Mann-Whitney $U = 31$, $p = .15$) or for uncommon targets (45% for the older group, range: 17–79%, versus 35% for the younger group, range: 13–71%; Mann-Whitney $U = 35$, $p = .27$). However, Figure 5 also shows that the pattern of responses was different for the two age groups: The younger children responded similarly to top- and underside-of-arm targets, but the older group showed comparably higher rates of target matches for uncommon gestures (approaching those produced to common gestures) in the top-of-arm trials, and a marked difference in responses to common and uncommon targets in the underside-of-arm trials (see the preceding section).

Children's matching errors: Common for uncommon. Figure 5 (bottom right panel) presents the percentages of trials on which common target gestures, uncommon target gestures, and any other incorrect responses were emitted as target mismatches to the Set 3 and Set 4 models. This figure shows that, in each target set and age group, children produced common target gestures as mismatches more frequently than either uncommon targets or all other gestures. Overall, across all Set 3 and Set 4 trials, common target gestures occurred as mismatches on 23% of trials (range: 2–42%) and uncommon targets on 13% of trials (range: 0–52%); this difference was statistically significant (Wilcoxon's $T = 50.5$, $p = .04$). Likewise, common target gestures occurred as mismatches much more frequently than all other gestures (8% of trials; range: 0–40%; Wilcoxon's $T = 19.5$, $p = .0001$).

Within common target mismatches, the children produced top-of-arm responses (T17 and T19) more often than underside-of-arm responses (T14 and T16; Wilcoxon's $T = 25$, $p = .0001$). The same pattern was recorded for uncommon target mismatches, where the children produced top-of-arm responses (T13 and T15) more often than underside-of-arm responses (T18 and T20; Wilcoxon's $T = 16.5$, $p = .04$). Most children in both groups also performed gestures other than targets from Set 3 and Set 4 as mismatches (and see Figure 6). First, touches to shoulder (Set 2 target gestures) were emitted on 2% of all Set 3 and Set 4 trials (range: 0–20%); 6 younger and 4 older children produced some of these responses. Second, various gestures in which hand(s) touched body parts other than lower- or middle-arm occurred on 6% of all trials (range: 0–33%): These mismatches mostly consisted of touches to incorrect arm locations (e.g., side of arm between wrist and elbow or upper arm) and of well-established matching responses, such as hugging, clapping, or touching eyes or head (B1 and B3); 7 younger and 7 older children emitted some such responses.

Summary: Set 3 and Set 4

All children produced target matches more frequently in response to commonly trained target models than to similar gestures that had no known prior training history in the nursery the children attended; they also performed common gestures as target mismatches more frequently than uncommon gestures, or any other responses. Across the middle- and lower-arm target trials of Set 3 and Set 4, children emitted top-of-arm responses more frequently than underside-of-arm responses, especially as target mismatches.

Verbal Responses

Children's verbal responses to experimenter's prompts, modeling, and their own gestural responses, in all experimental trials, are presented in Table 4. The contingencies did not promote such responses—the experimenter ignored all children's comments in testing; nonetheless, 5 children (younger Cal and Ette; older Cai, Stef, and Fion) named gestures and/or body parts on some of their trials; 8 children (younger Cal and Isa; older Tom, Fion, Stef, Ren, Cara, and Cai) produced other verbal responses, such as matching prompts directed at the experimenter. Although the data show that some children named some of the body parts that the experimenter touched in the course of modeling the target behaviors, it remains to be determined whether naming

Table 4

A summary of children's verbal responses to experimenter's prompts, modeling, and their own gestural responses.

Behavior	Description	Child
Naming of gestures and/or body parts	The child named ''head'' in response to a B3 model, and ''hands and tummy'' to a B2 model; in both cases he then responded correctly.	Cal
	The child said ''snap-snap!'' in response to a wrist-touch (T17) model, then proceeded to mime a crocodile jaw snap by placing a hand on top of backhand and opening and closing (this was coded as a correct response according to criteria but was clearly a different action to that modeled by the experimenter).	Cai
	The child observed modeling of a head-touch (B3) and said ''shoulders'' (the next line in the ''heads & shoulders'' naming and matching game); she then responded by touching her shoulders (T5). Also named ''tummy'' on several trials when B2 gesture was modeled and said ''peek-a-boo'' when B1 was modeled (and then responded correctly in all cases).	Fion
	On several occasions, the child said, ''my shoulders'' in response to modeling of unimanual shoulder touches (T8 & T9), then responded incorrectly by bimanual touches to shoulders (T11). Also on several occasions, she responded with ''boink!'' to modeling of bimanual ear touches, then responded by pulling both her ears (this was coded as correct according to our criteria, but was clearly a different action to that modeled by the experimenter).	Ette
	On two occasions, the child responded correctly to an unimanual ear touch (T1 & T3), immediately said, ''two ears'' and changed his response to an incorrect bimanual ear touch (T5).	Stef
Other verbal responses	The child said, ''No'' and did not respond immediately to a model on one or more occasions.	Cal, Tom, Fion, Isa
	The child said, ''I've done this already!'' as the second trial of a target was presented shortly after the first trial.	Ren
	The child said, ''I want to do this one!'' as she responded incorrectly to a target model.	Ren
	In response to a probe model, the child emitted a baseline gesture while saying to the experimenter, ''Can you do this?'' or, ''Do this!''	Fion, Stef
	The child repeated, ''Can you do…'' after the experimenter on several occasions (echoing the prompt).	Fion
	After responding incorrectly to a target gesture, the child said to the experimenter, ''Do what I do!'' or, ''Copy me!''	Cara, Cai
	On several trials where touches to wrist (T17) were modeled, the child commented on his shirt (e.g., ''My sleeve is open!'') while responding, or commented on the experimenter (e.g., ''I don't have a watch like that!'') before responding.	Cai

serves a self-instructional function that mediates matching performances in young children.

DISCUSSION

Young children's matching responses to 20 gestural models, each of which showed a hand touch to either one or two body locations, were investigated using the standard paradigm employed in behavior analysis to test for generalized imitation. As in previous studies reported in the developmental literature (Bekkering et al., 2000; Gleissner et al., 2000; Wohlschlager et al., 2003), which employed similar target gestures, the 3-year-olds made a large number of matching errors (40% of their responses were mismatches). The similarity in outcome across these developmental studies, that did not present target behaviors in the context of trained baseline matches but presented noncontingent reinforcers on all trials, and the present study, in which unreinforced target behavior trials were interspersed with intermittently reinforced baseline trials, suggests that a reinforcement-for-matching context does not result in better matching performances in 3-year-old children. This finding needs to be verified by conducting a within-study comparison of the two test paradigms. The present study extended the investigation of young children's matching performances by also presenting the target models to

2-year-olds and found even higher error rates (56% mismatches). Whereas Horne and Erjavec (2007) found no reliable evidence that infants produce matching responses to novel target behaviors in generalized imitation tests, the present study shows that children's matching performances improve somewhat late in the third year though they continue to produce matching errors in the majority of test trials. Matching accuracy continues to improve gradually throughout childhood, though other studies show that even older children continue to produce mismatched responses to some target behaviors, particularly those consisting of hand touches to body locations (Bekkering et al., 2000; Schofield, 1976).

Biases in Children's Responses During Imitation Tests

Mirror responses. In the estimates of matching responses described above, the difference between correct and mirror responses is ignored because children younger than 14 years of age (Bergès & Lézine, 1965; Wapner & Cirillo, 1968) frequently respond to a modeled gesture with its mirror counterpart. For example, when the experimenter faces the participant and models a right-hand touch to right ear this is most likely to evoke a left-hand touch to left ear—the observer has a strong tendency to respond in the same hemispace as the modeler. However, if mirror responding is subtracted from the matching responding reported above, these young children produced only 10% correct responding. This reliable observation that even responses with the correct topography tend to be mirror images of the behavior modeled is not explained by any of the cognitive accounts of imitation in young children. However, the tendency to mirror can be explained quite easily in a Skinnerian account of how the matching repertoire is established in the course of human development. When parents train their children to perform a new behavior this is usually conducted facing the child and when necessary moving the child's arm or leg through the target movement. After several trials of such ''putting through'' (Konorski & Miller, 1937) the child is encouraged to perform the behavior without manual guidance. Initially, the child responds *while* the adult is performing the target action, but with a slight lag, so as to produce a kind of ''visuomotor tethering'' of the child's behavior to that of the adult. Throughout such shaping procedures, the child usually performs the target behavior in the same hemispace as the parent. The prevalence of such training procedures is consistent with the dominance of mirror responding even in the 2-year-olds in the present study, a finding that extends the data on mirroring to children of this age; instances of mirroring have also been recorded in infant's responses to manual target gestures (Erjavec, 2002; Horne & Erjavec, 2007).

Hand biases. Both the 2-year-olds and the 3-year-olds tended to respond with their dominant hands when presented with unimanual models, as was the case in a study with 3- to 6-year-old children conducted by Bekkering et al. (2000). Consistent with a prevalence of right-handedness in the human population, most children tended to produce a right hand response whether a left hand model or a right hand model was presented and some children did so almost exclusively. Two children in each age group showed the converse bias in responding. A right hand response bias was also found for 3 out of 4 infants aged between 15 and 25 months (Erjavec, 2002, Experiment 2) which provides evidence that handedness biases occur early in human development.

Children in both age groups tended also to respond with two hands when a unimanual gesture was modeled, but did not show the converse tendency—there were very few single-handed responses to bimanual models. This response bias, also reported by Bekkering and colleagues, may originate in children's prior learning of lap games and nursery play routines, many of which incorporate bimanual responses such as clapping hands and touching head, shoulders, and toes. Whereas this error pattern is easily accommodated in a social learning perspective, it is not clear how the goal theory of Bekkering and colleagues could explain this tendency to respond with two hands instead of one.

Ipsilateral for contralateral errors. The strong tendency for young children and adult aphasia patients to respond to a contralateral model with an ipsilateral response is well documented (Bekkering et al., 2000; Benton, 1959; Gleissner et al., 2000; Gordon, 1923; Head, 1920, 1926; Kephart, 1971; Schofield, 1976; Wapner &

Cirillo, 1968). The same response bias was found in the present study in both the 2-year-olds and the 3-year-olds, for the contralateral ear touch models of Set 1 and the contralateral shoulder touch models of Set 2. The 2-year-olds produced significantly more such errors than the 3-year-olds, who in turn produced more errors than the 4- to 6-year-olds in Bekkering et al. (2000). Again, there appears to be a decreasing trend in such errors with development.

In terms of their goal theory, Bekkering et al. (2000) explain these ipsilateral responses to contralateral models as being due to young children's cognitive processing limitations. According to the theory, when a behavior is modeled, children identify which are the main goals and which are lesser goals of the behavior. For example, if an ear touch is modeled, the child infers that the main goal is to touch an ear—whether this goal is achieved by using one hand or two, an ipsilateral arm movement or a contralateral one, is less important because these components of the behavior are subsidiary goals. Consistent with the goal account, these authors report that young children produce more contralateral errors when the modeled gesture terminated on two dots on a table top than when there were no such markers on the table top. In a later study, Gleissner et al. (2000) demonstrated that children responded with 10% fewer errors when the behavior modeled was a contralateral response to a position near the ear as compared with one touching the ear. Thus, when the target behavior does not terminate on a specific object (a body part or two dots on a table top) the main goal becomes the cross-body movement, which the children produce more frequently. However, their 3-year-old participants still responded incorrectly on 20% of the contralateral trials in the ''near'' condition.

Common versus uncommon matching errors. Gleissner et al. (2000, p. 413) conclude that children's inferences about the goals of the target behaviors they see modeled are more important determinants of matching than is the ''raw familiarity with a gesture''. However, a trained matching account predicts that matching history is an important determinant of young children's matching performances. The child's social environment determines his or her exposure to particular target gestures in the course of matching games. In the present study, it was determined a priori which target gestures were commonly presented to all children in the nursery where this study was conducted. In the case of the Set 1 and Set 2 gestures, ipsilateral ear and shoulder touches, unimanual and bimanual, featured in the nursery games and were therefore predesignated as commonly trained target gestures whereas the contralateral counterparts were classified as uncommon because they did not feature in these games. Consistent with a trained matching account, the data show that commonly trained ear and shoulder touches were matched more frequently than their uncommon counterparts (see Figure 5, top left panel and top right panel). The same figure shows that common responses were also produced as mismatches to uncommon models more frequently than vice versa. However, because commonly-trained gestures are also ipsilateral, and uncommonly-trained gestures are contralateral, it is not possible to conclude that ipsilateral gestures are matched better because they already feature in the trained matching repertoires of many of the young participants, and that contralateral ones are matched poorly because they do not.

Therefore, Set 3 and Set 4 were devised to investigate the common versus uncommon dimension independently of the ipsilateral versus contralateral dimension in relation to young children's matching performances. All these gestures were contralateral touches to body parts but only half were commonly trained. In addition, half were ''easy'' gestures and half were ''difficult''. For the easy gestures, one arm was extended in a resting position on the table and the target part of that arm (either mid-arm—uncommon gesture, or wrist—common gesture) was touched with the other hand. For the difficult gestures, one arm was raised and bent so that the other hand could touch the target part of that arm (either elbow—common gesture, or wrist—uncommon gesture). ''Difficulty'' was therefore crossed with ''commonly trained'' over the eight target gestures. In this key test of the influence of the matching training environment, the children once again matched the commonly trained contralateral gestures significantly more often than uncommon contralateral gestures, at both ages, and regardless of whether the target gesture was easy or difficult to perform. First, the goal theory account does not predict that common gestures will be matched better than uncommon ones. Second, it predicts that children will neglect to raise their arm in the case of the difficult gestures because it is a lesser goal, but that they will touch the elbow or wrist accurately because that is the main goal. When shown the difficult (uncommon) arm raised, touch-to-wrist gesture the children did indeed produce instead the easy (common) arm down, touch-to-wrist gesture as often as, or even more often than, the difficult behavior modeled—an outcome consistent with goal theory. However, this neglect of the arm raise component did not happen in the case of the mid-arm (elbow) gestures. For these gestures the children raised their arms and touched their elbows when shown the difficult (but common) touch-to-elbow and produced instead very few (easy but uncommon) touches-to-elbow with arm in the lowered position as incorrect responses—an outcome consistent with a trained matching account but not with goal theory. Third, goal theory predicts that the children should not make elbow-for-wrist or wrist-for-elbow errors since the main goal is to touch the correct body location. However, over the 20 children, there were 20% such errors; wrist-for-elbow errors occurred as frequently as elbow-for-wrist errors. In addition, wrist-for-elbow errors were twice as frequent for 2-year-olds as for 3-year-olds, most probably because tapping the wrist is trained earlier in lap games than touching the elbow in body part naming games. All the main predictions of goal theory, which is the main cognitive developmental account of young children's imitation performances, are therefore contradicted by the present data.

The conditioned reinforcer hypothesis. Clearly, matching training is a major determinant of young children's responses to modeled target behaviors and children continue to produce mismatches during their early- to mid-childhood and perhaps beyond. How then do these data sit with the hypothesis put forward by Baer and Deguchi (1985) that young children may, as the trained matching repertoire develops, reach the point where they discriminate the parity between the modeled behavior and their own matching responses, whereupon the similarity acquires conditioned reinforcing properties? Although Horne and Erjavec (2007) found no evidence that parity served as a conditioned reinforcer for the matching responses of infants, this may be because they have small matching repertoires as compared with older children. However, neither the 2-year-olds nor the 3-year-olds in the present study showed systematic improvements in matching accuracy over trials. This suggests that when children of this age produce a response that, from an objective point of view, is a better match to the model than their previous responses, they are still not able to discriminate that increased similarity. The fact that matching errors persist in the performances of even older children suggests that the property of similarity is not a reliable determinant of young children's matching performances, or at best that it remains a very approximate determinant of their matching behavior.

This lack of evidence for an effect of parity on matching of manual responses differs from that for verbal behavior: Several studies have shown that matching accuracy of nonreinforced verbal target behaviors increases in a generalized imitation test paradigm (Brigham & Sherman, 1968; Lovaas, Berberich, Perloff, & Schaeffer, 1966; Schroeder & Baer, 1972). Brigham and Sherman found that normal preschool children matched nonreinforced Russian words with increased accuracy over test trials. However, in the case of verbal behavior, auditory feedback is likely to result in better discrimination of parity than can occur on manual behavior matching trials.

The Effects of Social Stimulus Enhancement and Naming on Matching Responses

Though they often did so with the incorrect trajectory of arm movement and/or two hands instead of one, the 2-year-olds and 3-year-olds in the present study frequently touched the correct body part. This finding can also be accommodated within a trained matching account which predicts that the endpoint of a behavior is likely to be more salient than the initial components of the behavior. From early in infancy, children learn to orient to a point gesture by looking at the object the caregiver is pointing to or touches (Butterworth & Cochran, 1980; Butterworth & Grover, 1988; Lempers, 1976; Messer, 1978; and see Horne & Lowe, 1996, p.193). This well-established history may explain why children respond more accurately to the body part touched rather than to the arm movement that preceded the touch. However, in verbal children, selective attention to Object rather than Action may also be the result of naming. According to the naming account of Horne and Lowe (1996), if the child names the part of the body touched by the modeler (e.g., ‘‘ear’’, ‘‘shoulder’’) this will evoke the listener behavior in the child of orienting to (e.g., touching) his or her ear. This self-instructional effect may facilitate the accuracy of touches to nameable body locations. The children were trained to touch their wrists and elbows in the course of nursery matching games, but the mid-arm location did not feature in those match-to-name games. It is interesting then that for the arm-down gestures in Set 3 and Set 4 that involved one location (‘‘wrist’’) previously named in matching games and one that was not (mid-arm), there were twice as many erroneous wrist touches to mid-arm models than vice versa. However, in the case of the raised-arm gestures for which both locations (‘‘wrist’’ versus ‘‘elbow’’) featured in nursery naming games there were few wrist-for-elbow errors; such elbow-for-wrist errors that did occur likely reflect the fact that raised elbow touches are commonly trained whereas raised wrist touches are not. The response competition that results from naming the target body part may also interfere with accurate imitation of the action that preceded the touch. However, if the child has no name for the endpoint of the gesture, as in the near body part conditions investigated in Gleissner et al., (2000), the child may instead name the action produced by the modeler and via name-based verbal governance thereby instruct him- or herself to perform the same action (Horne & Lowe, 1996, pp.199–208). In future studies, we will systematically investigate the role of action and body-part naming in the establishment of imitation in young children.

Conclusion

The Skinnerian account of imitation appears still to be the most complete account of how imitation is established in young children. We have shown that a competing cognitive developmental account does not explain straightforwardly the data in the present study in which gestures that are commonly trained in the course of matching games with young children are matched significantly more reliably than are gestures that have no such history. Interestingly, a recent study investigating mirror neuron activity also found that history determines whether or not mirror neuron activity occurs in the human brain when humans see others perform particular behaviors. Mirror neurons are situated in the premotor cortex, but—under certain conditions—they respond not only to execution of an action (as do other cells in this region), but also to passive observation of the same behaviour performed by others. It had been suggested that mirror neuron activity forms the neural basis for a variety of social responses, including imitation (Fadiga et al., 1995; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996), but the conditions under which this ‘‘special’’ supramodal neural system is established remain to be explored. Calvo-Merino, Glaser, Grèzes, Passingham, and Haggard (2005) used functional magnetic resonance imaging (fMRI) to study the relation between expertise in a particular physical skill and mirror neuron activity in adult human observers. Ballet dancers showed greater mirror neuron activity when watching a video of ballet dancing than when shown capoeira martial arts displays, whereas people skilled in the latter martial art but not ballet dancing showed the converse pattern. People who were neither trained in ballet nor the capoeira martial art showed much lower mirror neuron activity than those who had learned one or other of those skills and their brain activity was no different whether they viewed ballet dancing or martial arts displays. This study clearly shows an effect of history on mirror neuron activity when adult humans observe the behavior of others. As such, it is fully consistent with a trained matching account of imitation even in adult humans.

Together with our earlier study of imitation of manual gestures in infants (Horne & Erjavec, 2007) the present study provides developmental continuity in the detailed analysis of children’s matching repertoires over their next two years, identifying the characteristic errors made by young children during tests for generalized imitation and some of the social-environmental determinants of those errors.

REFERENCES

Anisfeld, M. (1996). Only tongue protrusion modeling is matched by neonates. *Developmental Review, 16,* 149–161.

Anisfeld, M. (2005). No compelling evidence to dispute Piaget's timetable of the development of representational imitation in infancy. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science* (Vol. 2, pp. 107–131). London: The MIT Press.

Baer, D. M., & Deguchi, H. (1985). Generalized imitation from a radical-behavioral view-point. In S. Reiss & R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 179–217). New York: Academic Press.

Baer, D. M., Peterson, R. F., & Sherman, J. A. (1967). The development of imitation by reinforcing behavioral similarity to a model. *Journal of the Experimental Analysis of Behavior, 10*, 405–416.

Bandura, A., & Barab, P. G. (1971). Conditions governing nonreinforced imitation. *Developmental Psychology, 5*, 244–255.

Bekkering, H., & Wohlschläger, A. (2002). Action perception and imitation: A tutorial. *Attention and Performance, 19*, 294–314.

Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53*, 153–164.

Benton, A. L. (1959). *Right-left discrimination and finger localization: Development and pathology.* New York: Hoeber-Harper.

Bergès, J., & Lézine, I. (1965). *The imitation of gestures: A technique for studying the body schema and praxis of children three to six years of age.* London: William Heinemann Medical.

Brigham, T. A., & Sherman, J. A. (1968). An experimental analysis of verbal imitation in preschool children. *Journal of Applied Behavior Analysis, 1*, 151–158.

Butterworth, G. E., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioural Development, 3*, 253–272.

Butterworth, G. E., & Grover, L. (1988). The origins of referential communication in human infancy. In L. Weiskrantz (Ed.), *Thought without language* (pp. 5–24). Oxford: Clarendon Press.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249.

Erjavec, M. (2002). *Determinants of gestural imitation in young children.* Unpublished doctoral dissertation, University of Wales: Bangor, UK.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology, 73*, 2608–2611.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593–609.

Garcia, E., Baer, D. M., & Firestone, I. (1971). The development of generalized imitation within topographical boundaries. *Journal of Applied Behavior Analysis, 4*, 101–112.

Gleissner, B., Meltzoff, A. N., & Bekkering, H. (2000). Children's coding of human action: Cognitive factors influencing imitation in 3-year-olds. *Developmental Science, 3*, 405–414.

Gordon, H. (1923). Hand and ear tests. *British Journal of Psychology, 13*, 283–300.

Griffiths, R. (1954). *The abilities of babies.* London: the Test Agency Limited.

Head, H. (1920). Aphasia and kindred disorders of speech. *Brain, 43*, 87–165.

Head, H. (1926). *Aphasia and kindred disorders of speech.* New York: MacMillan.

Horne, P. J., & Erjavec, M. (2007). Do infants show generalized imitation of gestures? *Journal of the Experimental Analysis of Behavior, 87*, 63–83.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*, 185–241.

Hurley, S., & Chater, N. (2005). Introduction: The importance of imitation. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation : From neuroscience to social science* (Vol. 2, pp. 1–52). London: The MIT Press.

Kaye, K. (1982). *The mental and social life of babies: How parents create persons.* Chicago: University of Chicago Press.

Kephart, N. C. (1971). *The slow learner in the classroom.* Columbus, OH: Charles Merrill.

Kokkinaki, T. (2003). A longitudinal, naturalistic, and cross-cultural study on emotions in early infant-parent imitative interactions. *British Journal of Developmental Psychology, 21*, 243–258.

Kokkinaki, T., & Vasdekis, V. G. S. (2003). A cross-cultural study on early vocal imitative phenomena in different relationships. *Journal of Reproductive and Infant Psychology, 21*, 85–101.

Konorski, J., & Miller, S. (1937). On two types of conditioned reflex. *Journal of General Psychology, 16*, 264–272.

Lempers, J. D. (1976). *Production of pointing, comprehension of pointing, and understanding of looking behavior in young children.* (Doctoral dissertation, University of Minnesota). *Dissertation Abstracts International, 37*, No. 6, p. 3524A.

Lovaas, O. I., Berberich, J. P., Perloff, B. F., & Schaeffer, B. (1966). Acquisition of imitative speech by schizophrenic children. *Science, 151*, 705–707.

Meltzoff, A. N. (1990). Foundations for developing a concept of self: The role of imitation in relating self to other and the value of social mirroring, social modeling, and self practice in infancy. In D. Cicchetti & M. Beeghly (Eds.), *The self in transition: Infancy to childhood* (pp. 139–164). Chicago: University of Chicago Press.

Meltzoff, A. N. (2005). Imitation and other minds: The ''like me'' hypothesis. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science* (Vol. 2, pp. 1–52). London: The MIT Press.

Meltzoff, A. N., & Moore, M. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*, 75–78.

Meltzoff, A. N., & Moore, M. (1983). Newborn infants imitate adult facial gestures. *Child Development, 54*, 702–709.

Messer, D. J. (1978). The integration of mothers' speech with joint play. *Child Development, 49*, 781–787.

Pawlby, S. (1977). Imitative interaction. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction.* London: Academic Press.

Poulson, C. L., Kymissis, E., Reeve, K. F., Andreatos, M., & Reeve, L. (1991). Generalized vocal imitation in infants. *Journal of Experimental Child Psychology, 51*, 267–279.

Poulson, C. L., Kyparissos, N., Andreatos, M., Kymissis, E., & Parnes, M. (2002). Generalized imitation within three response classes in typically developing infants. *Journal of Experimental Child Psychology, 81*, 341–357.

Schofield, W. N. (1976). Do children find movements which cross the body midline difficult? *Quarterly Journal of Experimental Psychology, 28*, 571–582.

Schroeder, G. L., & Baer, D. M. (1972). Effects of concurrent and serial training on generalized vocal imitation in retarded children. *Developmental Psychology, 6*, 293–301.

Sherman, J. A., Clark, H. B., & Kelly, K. K. (1977). Imitative behavior of preschool children: the effects of reinforcement, instructions, and response similarity. In B. C. Etzel, J. M. LeBlanc & D. M. Baer (Eds.), *New developments in behavioral research: Theory, method, and application. In honor of Sidney W. Bijou* (pp. 503–529). Hillsdale, NJ: Lawrence Erlbaum Associates.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Steinman, W. M. (1970). Generalized imitation and the discrimination hypothesis. *Journal of Experimental Child Psychology, 10*, 79–99.

Striano, T., & Liszkowski, U. (2005). Sensitivity to the context of facial expression in the still face at 3-, 6-, and 9-months of age. *Infant Behavior and Development, 28*, 10–19.

Want, S. C., & Harris, P. L. (2002). How do children ape? Applying concepts from the study of non-human primates to the developmental study of 'imitation' in children. *Developmental Science, 5*, 1–41.

Wapner, S., & Cirillo, L. (1968). Imitation of a model's hand movements: age changes in transposition of left-right relations. *Child Development, 39*, 887–894.

Williamson, R. A., & Markman, E. M. (2006). Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. *Developmental Psychology, 42*, 723–731.

Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: an instance of ideomotor principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 358*, 501–515.

Young, J. M., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1994). Generalized imitation and response class formation in children with autism. *Journal of Applied Behavior Analysis, 27*, 685–697.

Received: January 18, 2007
Final acceptance: November 29, 2007